



Corela

Cognition, représentation, langage

5-2 | 2007

Vol. 5, n° 2

La traduction automatique de phraséologismes pragmatiques : quelles représentations à travers la diversité formelle et culturelle ?

Aleksandra Dziadkiewicz



Édition électronique

URL : <http://journals.openedition.org/corela/383>

DOI : 10.4000/corela.383

ISSN : 1638-573X

Éditeur

Cercle linguistique du Centre et de l'Ouest - CerLICO

Référence électronique

Aleksandra Dziadkiewicz, « La traduction automatique de phraséologismes pragmatiques : quelles représentations à travers la diversité formelle et culturelle ? », *Corela* [En ligne], 5-2 | 2007, mis en ligne le 30 décembre 2007, consulté le 19 avril 2019. URL : <http://journals.openedition.org/corela/383> ; DOI : 10.4000/corela.383

Ce document a été généré automatiquement le 19 avril 2019.



Corela – cognition, représentation, langage est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

La traduction automatique de phraséologismes pragmatiques : quelles représentations à travers la diversité formelle et culturelle ?

Aleksandra Dziadkiewicz

Introduction

- ¹ Dans ce travail, nous présentons un modèle théorique de traduction automatique de phraséologismes pragmatiques, basé sur la constitution de représentations intermédiaires contenant des éléments abstraits, communs aux langues analysées. Le modèle a été conçu à l'origine pour la traduction du français vers le polonais, mais nous pensons qu'il peut convenir également à la traduction d'autres couples de langues. Forgé par Burger (1973, 1982), le terme de phraséologismes pragmatiques (PP) désigne un large groupe de formulations conventionnelles réalisant des actes de langage déterminés qui ne peuvent être décrits que dans un cadre pragmatique, c'est-à-dire faisant référence à la situation de leur énonciation. L'auteur a inclus parmi les PP des expressions fixes telles que *Guten Tag* (Bonjour) et *Guten Appetit* (Bon appétit), mais aussi des constructions lexico-grammaticales ouvertes comme *Würden Sie mir bitte den Zucker herüberreichen ?* (Me passeriez-vous le sucre, s'il vous plaît ?) ou *Hätten Sie die Freundlichkeit mich ins Theater zu begleiten ?* (Auriez-vous l'amabilité de m'accompagner au théâtre ?). Ce type de constructions n'a pas suscité jusqu'à présent un grand intérêt de la part des chercheurs en traitement automatique des langues qui privilégient, en général, des études portant sur le degré de figement des lexies complexes non compositionnelles telles que les noms composés et différents types de locutions (verbales, adjectivales, adverbiales, conjonctives...). On ne pourrait surestimer l'apport de ces travaux pour l'informatisation des données linguistiques et par là même pour la traduction automatique. Cependant, une traduction

automatique de qualité, et la traduction en général, nécessitent une prise en compte de toutes les structures préfabriquées de la langue, qu'elles soient compositionnelles ou non. Comme l'affirme Bourquin (1993 : 29),

il faut tenir compte de toute une tradition de formulations, rechercher tout ce qui, à force de se répéter, tend à se scléroser, à se figer, faire l'inventaire de stéréotypes discursifs de tous ordres. C'est alors seulement qu'on peut localiser les écarts de langue à langue, et construire les procédures pour les réduire.

- 2 Nous considérons que les phraséologismes pragmatiques font partie des stéréotypes discursifs et que, de ce fait, ils doivent être inventoriés et analysés en vue de traduction automatique.
- 3 Nous avons limité notre objet d'analyse à un groupe de PP sélectionnés et classés selon quelques actes de langage dits sociaux : le remerciement, l'excuse, la requête, la proposition, la salutation d'ouverture et la salutation de fermeture. Ce choix a été dicté par le fait que ces actes semblent être universels ou quasi universels, tout en prenant des formes particulièrement déterminées par les facteurs socioculturels. A titre d'exemple, une requête formulée au moyen d'une structure interrogative avec le verbe *pouvoir*, couramment utilisée dans les langues occidentales, serait mal perçue en Thaïlande (Kerbrat-Orecchioni, 1998 : 42). De même, la stratégie sémantique consistant à « déguiser » une requête en remerciement (*merci de faire...*), bien connue des Français, risque de ne pas être comprise par les Polonais. Les critères syntaxiques ou purement sémantiques ne suffisent donc pas à obtenir une bonne traduction : celle-ci gagne en qualité grâce à l'introduction d'informations d'ordre pragmatique et notamment de celles qui spécifient la nature de l'acte de langage réalisé par chaque PP. La détermination de ces informations a été un point crucial pour l'élaboration de notre modèle de traduction, dont les composantes seront présentées dans le corps de l'article.

1. Diversité formelle et culturelle des PP

- 4 Tout d'abord, nous passerons en revue plusieurs exemples de divergences culturelles et linguistiques à travers les réalisations d'actes de langage : nous verrons ainsi comment les stratégies sémantiques directes et indirectes varient d'une langue à l'autre et à quel point la prise en compte de ces variations est importante pour la traduction.

1.1. Divergences de formulation du même acte de langage¹

- 5 Prenons l'exemple des formules de salutation : elles varient d'une société à l'autre et, pour un locuteur non natif, elles sont parfois difficilement décodables en tant que salutations. Ainsi, au Vietnam, la forme des questions rituelles de salutation dépend de la situation de l'énonciation : *qu'est-ce que tu fais là ? Tu achètes du riz, n'est-ce pas ?* (au marché), *Tu vas acheter du riz ? / Tu vas au marché ?* (dans la rue), *Tu as mangé du riz ?* (à la cafétéria). Ces formules servent conventionnellement à assurer l'ouverture de la conversation et ne demandent qu'une réponse succincte.
- 6 Chez les Touaregs, ces questions portent sur la situation météorologique : *Qu'est-ce que tu as fait (avec) la pluie / la chaleur / le vent chaud / la sécheresse ?* Citons aussi l'exemple d'une formule de salutation très métaphorique ou même poétique, qui serait probablement impossible à décoder pour un Européen et qui est pourtant courante chez les Dowayo du Nord-Cameroun : *Le ciel est-il clair pour toi ? – Le ciel est clair pour moi, et pour toi ?* Enfin,

citons encore la salutation de fermeture utilisée chez les Eskimos qui prend la forme d'un constat : *Je m'en vais. – Tu t'en vas.*

1.2. Divergences d'interprétation du même contenu propositionnel

- 7 Les divergences d'interprétation que les usagers de différentes langues attribuent à un énoncé peuvent donner lieu à des malentendus pragmatiques. Afin de tester ces divergences culturelles et linguistiques, Fonagy (1982) avait proposé à un groupe de locuteurs français et hongrois d'interpréter des énoncés habituellement liés à des situations précises. Les résultats de cette enquête ont montré que l'interprétation dépend dans une grande mesure de la façon dont on décode conventionnellement certaines formules dans chacune des deux langues. Par exemple, la formule *Il ne fallait pas* a été associée à un remerciement par dix-sept Français sur vingt, alors que les sujets hongrois y ont vu un reproche réel formulé suite à une action indésirable faite par l'allocutaire.
- 8 La traduction directe du français vers le hongrois d'un énoncé comme *il ne fallait pas* ne prendrait pas en compte la tradition discursive de chacune des langues et, par conséquent, donnerait lieu à une fausse interprétation. C'est pourquoi il est intéressant de recenser tous les stéréotypes discursifs et d'assurer le passage d'une langue à l'autre en utilisant des concepts abstraits d'actes de langage.

1.3. Divergences et traduction

- 9 Certaines divergences culturelles sont tellement importantes qu'il serait extrêmement difficile, voire impossible, de les traiter en traduction automatique. C'est une des raisons pour lesquelles nous avons choisi comme objet d'étude les actes de langage réputés pour être sinon universaux, du moins largement présents dans la plupart des langues. Il est vrai que, même à l'intérieur de ce groupe de PP, certains sont spécifiques à une langue donnée et correspondent à des situations bien précises : le turc par exemple dispose d'une formule, *Ellerin dert görmesin* (*que tes mains ne connaissent jamais de soucis*), utilisée pour remercier d'une action faite avec les mains (Tannen, Comert Oztek, 1981 : 52-54). Pour utiliser correctement les formules de remerciement en turc, il faut donc connaître la nature de l'objet du remerciement. Or, celle-ci n'est pas explicitée dans les formules françaises : par conséquent, en traduction du français vers le turc, on ne saurait dire quelle formule turque conviendrait le mieux. Par conséquent, notre modèle de traduction fait appel à des critères plus généraux qui permettent de sélectionner une traduction approchante, c'est-à-dire un PP en langue cible qui ait le plus grand nombre de caractéristiques pragmatiques en commun avec le PP de la langue source.
- 10 D'autres problèmes peuvent apparaître lors de la traduction des PP. Certaines langues, comme le hongrois, font la distinction entre les salutations échangées par les adultes et les salutations adressées par un enfant à un adulte. Ainsi, même si la formule hongroise *Jónapot, uram* est sémantiquement équivalente à *Bonjour, Monsieur*, il est inimaginable qu'elle soit employée par un enfant s'adressant à un adulte : c'est une formule spécifique à cette situation, *Kezicsókolom* (*Je vous embrasse les mains*), qui s'impose (Fonagy, 1982 : 82). Cette différence n'étant pas formellement marquée en français, il est impossible de trouver la traduction hongroise pragmatiquement correspondante sans connaître l'âge des interlocuteurs. Une traduction entièrement automatique ne pourrait tenir compte de ces données puisqu'elles ne sont pas encodées dans le texte.

- 11 Il arrive aussi que là où une langue donnée dispose d'un énoncé propre à une situation, d'autres langues ne proposent aucune solution verbale : en Grèce et en Turquie, l'achat d'un nouveau vêtement est normalement commenté par un acte de vœu, *Me gheia* (avec *santé*) en grec et *Güle güle giy* (portez-le en riant, dans la joie) en turc (Tannen, Comert Oztek, 1981 : 52-54), alors qu'il n'existe aucune formule équivalente en français. Les Turcs ont l'habitude d'émettre des vœux de bonne santé, *Sihhatler olsun* (Que cela soit bon pour ta santé), à quelqu'un qui vient de prendre un bain ou de se couper les cheveux (*ibid.*), tandis que cet usage est inconnu dans les cultures occidentales.
- 12 Enfin, précisons encore une dernière chose : comment traiter les divergences concernant l'usage des formules. Comme le remarque Coulmas :
- One of the central problems in contrastive analysis is the relation between form and function in language. If we know *how* to say, *I'm sorry*, in another language, we still don't know *when* and to *whom* we should say it according to the norms of interaction of the respective community.
(Coulmas, 1981: 69)
- 13 A la lumière de la distinction entre la forme et la fonction (ou, si l'on veut, l'usage que l'on fait de cette forme), notre analyse doit se baser sur la forme : il est vrai, en effet, que la traduction automatique ne saurait probablement fournir de solutions concernant les règles de savoir-vivre. Autrement dit, si dans le texte source on note l'absence d'une formule de politesse, elle ne va pas être générée dans le texte cible. De même, une formule employée dans le texte source, qui dans certaines conditions pourrait paraître impolie aux usagers de la langue cible (par exemple, dans certaines sociétés le remerciement venant du client est mal perçu dans les relations commerciales (Kerbrat-Orecchioni, 2001 : 183)), sera malgré tout traduite dans le texte cible.

2. Reconnaissance formelle des PP

- 14 L'ensemble des unités analysées présente une grande hétérogénéité formelle : les PP peuvent être des mots-phrases (*enchanté* ; *amicalement*), des locutions-phrases (*s'il vous plaît*), des formules polylexicales (*je ne saurais jamais assez vous remercier*) ou des constructions lexico-syntaxiques (*pourriez-vous + infinitif + ?*)². Vu cette grande disparité, notre première tâche a été de nature descriptive : il s'agissait de cerner les caractéristiques et le fonctionnement dans le discours des unités analysées afin de permettre à la machine de les reconnaître au niveau formel. Nous avons examiné les variantes des PP étudiés, leurs propriétés transformationnelles, les possibilités de substitution, d'insertion et la manière dont ils se rattachent au reste de la chaîne. Les conditions formelles du fonctionnement de ces unités en tant que phraséologismes pragmatiques ont été encodées sous forme de règles de reconnaissance, inspirées de la syntaxe de l'analyseur morpho-syntaxique Labelgram, conçu et développé au Centre Tesnière, à l'Université de Franche-Comté. Ces règles sont composées d'éléments lexicaux spécifiques et de leurs variantes, ainsi que d'éléments grammaticaux plus généraux (parties du discours, reconnues par l'analyseur Labelgram), ce qui permet de reconnaître les PP admettant plusieurs transformations ou insertions d'éléments libres. Par exemple, l'étiquette *être* peut représenter dans une règle toutes les formes conjuguées du verbe *être*, tandis que les étiquettes *verbe* ou *adjectif* représentent respectivement tous les verbes ou tous les adjectifs de la langue³.

- 15 Nous voudrions développer ici un point important de la description formelle des PP : l'opposition établie entre la partie relationnelle et la partie informative. La partie relationnelle de l'énoncé véhicule les intentions du locuteur et définit sa position à l'égard de l'allocutaire, tandis que la partie informative apporte des éléments spécifiques, indépendants du PP même. Les proportions entre les deux parties peuvent varier selon le type de PP. Dans les formules de salutation d'ouverture et de clôture ou dans des formules concises d'excuse et de remerciement, la partie informative est pratiquement inexistante et la forme du phraséologisme reste fixe, on parlera alors de phraséologismes fermés :
- 16 *Salut ! Bonne journée ! Au plaisir ! Pardon ! Merci !*
- 17 D'autres PP, que nous appellerons ouverts, contiennent une partie relationnelle, représentée par les éléments constants de chaque expression, et une partie informative qui correspond à des éléments variables ou, autrement dit, à des « cases vides ». C'est le cas, par exemple, des formules de requête, où le locuteur est obligé de préciser l'action demandée :
- 18 *Merci de bien vouloir / nous retourner ce questionnaire.*
- 19 partie relationnelle partie informative
- 20 La partie relationnelle peut contenir elle aussi des éléments variables dépendant du contexte de l'énonciation, il s'agit des marques grammaticales du locuteur et de l'allocutaire : les pronoms personnels, les formes verbales, les termes d'adresse et les adjectifs possessifs. Ainsi, dans les formules épistolaires de clôture, les éléments variables sont sélectionnés en fonction des propriétés de chacun des participants :
- 21 *Je/nous vous prie/prions d'agréer, Monsieur/Madame, l'expression de mes/nos sentiments distingués.*
- 22 L'opposition entre la partie relationnelle et la partie informative de l'énoncé s'avère utile à la définition des éléments susceptibles de recevoir une description abstraite commune à plusieurs langues. On les retrouve dans la partie relationnelle, autant du côté des éléments constants que des éléments variables. Les premiers peuvent être étiquetés selon le type d'acte de langage exprimé ainsi que selon d'autres critères que nous exposerons plus bas : l'emploi (général vs spécifiques), le degré de politesse et leur stratégie sémantique. Les seconds véhiculent des informations sur le nombre, le sexe et le type de relation entre les interlocuteurs : ces propriétés existent dans le monde réel et doivent être considérées comme universelles, cependant elles trouvent leur reflet dans les langues dans des mesures différentes.

3. Constitution d'étiquettes sémantico-pragmatiques

- 23 Les paramètres qui rentrent dans les étiquettes de chaque phraséologisme relèvent de la sémantique (registre, stratégie sémantique) et de la pragmatique (type d'acte de langage, emploi, registre, degré de politesse, fréquence). Ils servent à décrire et à classer les PP aussi bien de la langue source que ceux de la langue cible.

3.1. Type d'acte de langage

- 24 Le type d'acte de langage est à la clef de notre classement et de notre modèle de traduction : il est l'élément principal de l'étiquette sémantico-pragmatique de tous les PP,

qui doit servir de pivot lors de la traduction de ceux-ci vers d'autres langues. La détermination du type d'acte de langage est donc indispensable pour que notre modèle fonctionne. Or, certaines structures posent des problèmes d'interprétation puisqu'elles peuvent exprimer plusieurs valeurs illocutoires. C'est le cas, par exemple, des constructions interrogatives avec le verbe *vouloir* qui fonctionnent soit comme une simple question (*Quel test voulez-vous effectuer ? Veux-tu être riche ? Songe à épargner autant qu'à gagner.*), soit comme un acte de proposition (*Veux-tu profiter de cette petite interview pour adresser un message à la communauté ? Veux-tu jouer avec moi ?*) ou de requête (*Veux-tu être plus précis dans tes explications ? Veux-tu m'expliquer la problématique du bac ?*). La résolution de ce problème implique une analyse contextuelle et la construction des algorithmes de désambiguïsation : cette question est abordée plus en détail dans Dziadkiewicz (2007b).

3.2. Emploi

- 25 Ce critère décrit les conditions d'emploi du phraséologisme. Certains PP sont propres à la langue écrite : c'est le cas notamment des formules épistolaires qui n'apparaissent jamais à l'oral avec la même fonction, comme par exemple les formules finales *Amitiés* ou *Cordialement*. Certes, les mots *amitiés* et *cordialement* fonctionnent également dans la langue orale, mais jamais en tant que PP et jamais en tant que salutations de clôture. Généralement, ils font alors partie de suites plus longues et il est rare qu'ils forment un énoncé à eux seuls⁴.
- 26 Par conséquent, nous distinguons les PP propres à la langue écrite de ceux dont l'emploi n'est pas spécifié. Nous parlerons donc d'emploi écrit et d'emploi général. Plusieurs types d'emplois spécifiques ont été également introduits dans la description des PP de salutation d'ouverture et de fermeture (par exemple l'emploi propre à la situation où le locuteur accueille l'allocutaire, pour les formules *soyez bienvenu*, *nous vous souhaitons la bienvenue*).

3.3. Registre

- 27 Le registre (ou niveau de langue) se manifeste dans les PP avant tout par certains choix lexicaux et syntaxiques. Nous avons adopté une tripartition en registres soutenu, standard et familier. Comme on pouvait s'y attendre, ce sont les PP du registre standard qui sont les plus nombreux : ils peuvent être employés aussi bien par des personnes qui se tutoient que par celles qui se vouvoient. Les PP du registre familier sont réservés aux interlocuteurs proches l'un de l'autre (le vouvoiement de politesse en est exclu), tandis que les PP du registre soutenu n'apparaissent que dans certaines conditions, lorsque les relations entre les interlocuteurs sont hiérarchiques et entraînent l'emploi des termes d'adresse de politesse. Voici des exemples de PP de remerciement appartenant aux trois registres : standard *nous tenons à vous dire merci*, soutenu *Je vous présente l'expression de ma gratitude* et familier *c'est super gentil (de ta part/à toi)*.

3.4. Degré de politesse

- 28 Une partie importante des PP est marquée par les principes de politesse. Dans certains groupes de PP, on peut observer une graduation à partir des formes neutres jusqu'aux formes les plus élaborées et les plus polies. Par exemple, le degré de politesse des

formules de remerciement augmente lorsqu'elles sont accompagnées de collocatifs intensifieurs (*merci ; merci beaucoup ; je vous dois un grand merci ; je te dis un énorme merci ; je te dis de tout mon cœur merci ; je te dis un merci des plus profonds...*). Les formules de requête, à leur tour, sont généralement plus polies en présence de marques d'atténuation (ajout de la formule *s'il te/vous plaît*, utilisation de constructions interrogatives avec *pouvoir* et *vouloir*, emploi du conditionnel et de l'imparfait : *je voulais vous demander de..., pourrais-tu faire...?* utilisation de formulations indirectes : *cela t'ennuierait de... ? merci de faire...je vous serais reconnaissant de faire...veuillez prendre la peine de faire...*).

- 29 L'établissement des équivalences entre différentes langues est très difficile lorsqu'on parle des règles de politesse. En effet, il n'existe pas une échelle de politesse unique pour toutes les langues et cultures : une langue A peut avoir cinq degrés de politesse possibles dans la formulation de la requête, tandis que la langue B en aura huit et la langue C n'en aura que trois. Dans une telle configuration, si l'on traduit de la langue C vers la langue B, on n'exploite pas tout le potentiel de la langue cible, puisqu'on est limité à choisir des formules relativement neutres. Dans le cas inverse, quand on traduit de la langue B vers la langue C, on est amené à opérer certaines réductions du sens, puisque la langue cible ne dispose pas de moyens d'expression aussi riches que la langue source. Dans notre modèle, dans le cas où la langue cible manquerait de degré équivalent au degré de politesse du PP de la langue source, c'est le degré le plus élevé de l'échelle de la langue cible qui est proposé en équivalence.

3.5. Stratégie sémantique

- 30 La stratégie sémantique d'un PP correspond grosso modo à son sens littéral. Un acte de langage peut être réalisé par une multitude de stratégies différentes. Par exemple, l'acte d'excuse en anglais « se décline » selon neuf stratégies sémantiques, recensées par Fraser (1981 : 263). Citons-en quelques unes :

- 31 • **Strategy 1:** Announcing that you are apologizing
 32 “I (hereby) apologize for...”
 33 • **Strategy 2:** Stating one's obligation to apologize
 34 “I must apologize for...”
 35 • **Strategy 3:** Offering to apologize
 36 “I (hereby) offer my apology for...”
 37 “I would like to offer my apology to you for...”
 38 • **Strategy 4:** Requesting the hearer accept an apology
 39 “Please accept my apology for...”
 40 “Let me apologize for...”
 41 “I would appreciate it if you would accept my apology for...”
 42 • **Strategy 5:** Expressing regret for the offence
 43 “I'm (truly / very / so / terribly) sorry for...”
 44 “I (truly / very much / so...) regret that I...”

- 45 L'utilisation de stratégies indirectes propres à l'anglais a été longtemps considérée par des chercheurs anglophones comme universelle (commune à toutes les langues), comme en témoigne l'affirmation suivante :

Every language makes available the same set of strategies - semantic formulas - for performing a given speech act. (...) if one can request, for example, in one language by asking the hearer about his ability to do the act (*Can you do that?*), by expressing one's desire for the hearer to do the act (*I'd really appreciate if you'd do that*), (...) then these same semantic formulas - strategies - are available to the speakers of every other language.

(Fraser, Rintell, Walters, 1980 : 78-79)

- 46 Cette prise de position a été par la suite remise en question et qualifiée d'ethnocentrique (Wierzbicka, 2003). Dans la première section de cet article, nous avons montré que les stratégies utilisées conventionnellement pour exprimer le même acte peuvent varier considérablement d'une langue à l'autre.

3.6. Fréquence

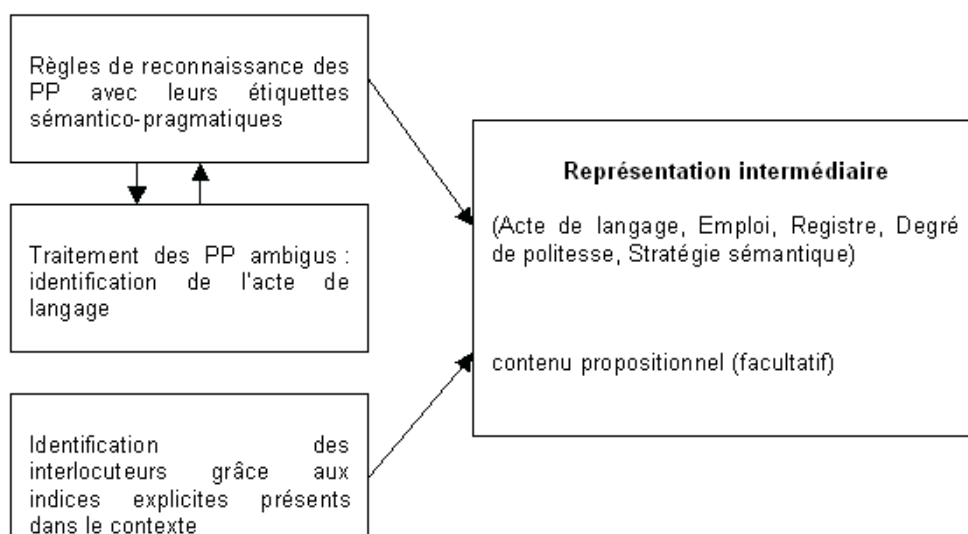
- 47 Le critère de fréquence est utile, lorsque plusieurs PP de la langue cible ont exactement la même étiquette sémantico-pragmatique. Dans ce cas-là, il faut en choisir un qui soit le plus représentatif de sa classe et lui assigner l'étiquette de PP canonique. Ce choix est déterminé précisément par la fréquence d'emploi de différents PP.

4. Construction de représentations intermédiaires

- 48 Dans notre modèle de traduction, les modules de reconnaissance et de désambiguïsation aboutissent à une représentation intermédiaire abstraite. Rappelons les grands volets de l'analyse des PP que nous avons évoqués jusqu'à présent :

- le phraséologisme est reconnu dans le texte grâce à la consultation des règles de reconnaissance du dictionnaire des PP source ;
- si ce phraséologisme correspond à deux actes de langage différents (il apparaît dans deux entrées dans le dictionnaire des PP), il subit le traitement adapté : soit l'ambiguïté est levée grâce à un algorithme, soit une solution par défaut est proposée (Dziadkiewicz, 2007b) ;
- l'étiquette sémantico-pragmatique du phraséologisme rentre dans sa représentation intermédiaire, elle est composée de cinq paramètres qui correspondent à la partie relationnelle du PP (Acte de langage, Emploi, Registre, Degré de politesse, Stratégie sémantique) ;
- le module d'identification des interlocuteurs⁵ participants à l'énonciation ajoute à la représentation intermédiaire du PP les paramètres définissant le locuteur (+/-sg, +/-masc.) et l'allocutaire (+/-sg, +/- masc., +/-poli). Ces paramètres, obtenus à partir des indices explicites présents dans le texte source, complètent la partie relationnelle du PP ;
- s'il s'agit d'un phraséologisme ouvert (composé d'une partie relationnelle et d'une partie informative), la représentation intermédiaire est également complétée par le contenu propositionnel de la partie informative : celui-ci doit être traduit au moyen d'un dictionnaire bilingue.

- 49 Schématiquement, ces modules peuvent être représentés de la façon suivante :



- 50 A partir de l'étiquette sémantico-pragmatique contenue dans la représentation intermédiaire, l'algorithme de traduction pourra proposer un PP cible avec l'étiquette sémantico-pragmatique identique ou la plus proche. Ce PP cible, sous forme de règle figurant dans le dictionnaire, peut générer plusieurs réalisations au niveau de la forme : sa forme finale, celle qui apparaîtra dans le texte cible, dépend en effet des paramètres des interlocuteurs ainsi que de l'éventuel contenu propositionnel, et notamment de la façon dont il se rattache au PP. A cette dernière étape de la traduction, celle de génération, les critères formels (morpho-syntaxiques) et stylistiques propres à la langue cible rentrent en jeu.

5. Construction d'un algorithme de traduction

- 51 Si l'on applique les démarches présentées dans les sections précédentes à deux langues, à l'issue de l'analyse on dispose de dictionnaires en langue source et en langue cible, contenant les PP avec leurs règles de reconnaissance et leurs étiquettes sémantico-pragmatiques, qui sont intégrées dans des représentations intermédiaires. On peut alors envisager d'établir des équivalences d'une langue à l'autre. Cette opération consiste à retrouver dans le dictionnaire de la langue cible un PP avec la même étiquette sémantico-pragmatique que le PP de la langue source. Des dictionnaires parfaitement symétriques, où à chaque PP de la langue source correspond un PP de la langue cible, n'existent pas, du fait de la grande diversité des langues : le nombre et le niveau de spécialisation des PP varient d'une langue à l'autre. Il est vrai que de nombreux PP dans une langue donnée trouvent facilement leurs équivalents exacts dans d'autres langues : même s'ils sont parfois bien différents au niveau de la forme, ils remplissent les mêmes conditions d'emploi, appartiennent au même registre, ont le même degré de politesse et exploitent la même stratégie sémantique. La traduction passe alors par l'association des PP source et cible ayant les étiquettes identiques.
- 52 Cependant, il peut y avoir aussi des divergences dans la constitution des étiquettes. Par exemple, un PP source peut avoir le même type d'emploi, même registre et le même degré de politesse qu'un PP cible, mais être construit autour d'une autre stratégie sémantique. Ou bien, un PP cible peut avoir le même type d'emploi, degré de politesse et stratégie sémantique qu'un PP source, mais appartenir à un autre registre. Lorsque de telles

divergences ont lieu, il convient de proposer tout de même un équivalent, sinon exact, du moins approchant. Nous essayons ainsi d'imiter le travail du traducteur humain qui, faute de trouver un équivalent exact, cherche à proposer une traduction approchante. Il paraît que plus un PP cible a de critères en commun avec le PP source, mieux il est adapté à être proposé comme équivalent, le cas le plus extrême étant celui où le seul trait commun est l'appartenance du PP source et du PP cible au même type d'acte de langage.

- 53 Afin de favoriser le choix des PP les plus standard en l'absence d'équivalence exacte, nous avons proposé les solutions suivantes :
- si parmi les PP cible il n'y a pas de candidat avec le même type d'emploi, chercher parmi les PP avec un emploi général,
 - si parmi les PP cible il n'y a pas de candidat avec le même registre, chercher parmi les PP appartenant au registre standard,
 - si parmi les PP cible il n'y a pas de candidat avec le degré de politesse équivalent, choisir le PP avec le degré le plus élevé,
 - si parmi les PP cible il n'y a pas de candidat avec la même stratégie sémantique, choisir le PP avec la stratégie sémantique canonique (la plus représentative du registre et degré de politesse en question).
- 54 Ces solutions doivent être appliquées tout en sachant que l'équivalence au niveau du registre, du type d'emploi et l'équivalence des degrés de politesse ont plus d'importance que l'équivalence des stratégies sémantiques. En effet, notre modèle de traduction relègue au second plan l'équivalence au niveau de la composante purement sémantique, en favorisant les critères d'ordre pragmatique. L'équivalence au niveau de la composante purement formelle est à plus forte raison écartée à ce stade.
- 55 Sans présenter en détail l'algorithme de recherche d'équivalence, précisons qu'il contient 15 conditions et que quatre instructions suffisent pour aboutir à l'un des 16 opérateurs.

6. Exemple du fonctionnement

- 56 Notre modèle théorique doit être en principe applicable à la traduction du français vers n'importe quelle langue. Cependant, nous nous servirons ici de l'exemple de traduction du français vers le polonais (décrit dans Dziadkiewicz, 2007a) pour illustrer le fonctionnement du modèle.
- 57 Nous citerons un texte authentique à traduire en y soulignant le PP. Nous le traduirons d'abord comme si c'était une simple suite libre, en imitant ainsi le traitement d'un traducteur automatique type. Ensuite, nous expliquerons toutes les étapes du traitement proposé par notre modèle, appliqué à cet exemple concret. Enfin, nous comparerons les deux traductions obtenues et présenterons les avantages de notre méthode de traduction.
- 58 Exemple d'un blog Internet :
- Dis-moi Lili, est-ce que ça te dérangerait de faire de la pub pour mon forum ? Je te demande ça parce qu'on est 7 dessus pour l'instant et ce serait bien que de bonnes âmes viennent nous rejoindre...
- <http://creationsdelili.canalblog.com/archives/2006/01/17/1238475.html>

6.1. Reconnaissance du PP dans le texte

- 59 Un PP est repéré grâce à une règle de reconnaissance, spécifiant ses éléments fixes et variables :
- 60 *est-ce que ça te dérangerait de faire de la pub pour mon forum ?*
- 61 Traduction littérale : *czy to by Ci przeszkadzało zrobić reklamę dla mojego forum ?*
- 62 Explication mot à mot de la phrase polonaise : particule interrogative – *cela* – marque du conditionnel – *te* (cas datif) – *dérangerait* – *faire* (forme perfective) – *pub* (cas accusatif) – *pour* – *mon* (cas génitif) – *forum* (cas génitif)
- 63 Cet énoncé serait compris par un locuteur polonais comme : *est-ce que cela (ce fait) te rendrait difficile l'action de faire de la pub pour mon forum ?* où *cela* serait obligatoirement considéré comme une reprise anaphorique d'un élément antérieur du texte.

6.2. Identification des interlocuteurs

- 64 Il convient de déterminer les paramètres de *je* et de *tu*. La première personne du singulier désigne le locuteur unique du sexe masculin ou féminin. Pour déterminer le paramètre +/-masc., l'analyseur vérifie si dans le texte à traduire, il y a des phrases où *je* est suivi d'un verbe copule et d'un adjectif ou participe au féminin. La réponse est négative et le locuteur prend les paramètres par défaut (+masc., +sg.)
- 65 La deuxième personne du singulier désigne l'allocutaire unique du sexe masculin ou féminin. Pour déterminer son paramètre +/-masc., l'analyseur vérifie si dans le texte à traduire, il y a des termes d'adresse nominaux ou prénoms en apostrophe. Dans la première phrase, *Lili* est un prénom féminin, le *tu* hérite du genre de ce substantif. Les paramètres de l'allocutaire sont donc (-masc., +sg.)

6.3. Constitution de la représentation intermédiaire

- 66 La représentation intermédiaire contient l'étiquette sémantico-pragmatique du PP à traduire (Acte : requête, Registre : standard, Emploi : général, Degré de politesse : 3, Stratégie sémantique : question sur l'objection de l'allocutaire), les propriétés des interlocuteurs (locuteur (+masc., +sg.), l'allocutaire (-masc., +sg.)) et le contenu propositionnel (CP : *faire de la pub pour mon forum ?*).

6.4. Algorithme de traduction

- 67 Puisqu'il n'y a pas parmi les PP polonais de PP ayant la même stratégie sémantique, l'algorithme aboutit à un PP canonique ayant les mêmes emploi, registre et degré de politesse :
- 68 *czy mógłbyś/mogłabyś/moglibyście/mogłybyście zrobić...?*
- 69 *czy mógłby/mogłaby/mogliby/mogłyby Pan/Pani/Panowie/Panie/Państwo zrobić...?*
- 70 Les différentes possibilités de réalisation formelle de ce PP, séparées par '/', dépendent des paramètres de l'allocutaire et correspondent à différentes combinaisons de traits +/-masc., +/-sg., +/-poli. En l'expliquant mot à mot, on obtient donc :

- 71 Particule interrogative – *pourrais* 2^{ème} personne(+masc., +sg., -poli) / (-masc., +sg., -poli) / (+masc., -sg., -poli) / (-masc., -sg., -poli) – *faire* (forme perfective)
- 72 Particule interrogative – *pourrait* 3^{ème} personne (+masc., +sg., +poli) / (-masc., +sg., +poli) / (+masc., -sg., +poli) / (-masc., -sg., +poli) – *Monsieur / Madame / Messieurs / Mesdames / Madame, Monsieur* – *faire* (forme perfective)

6.5. Génération du PP cible

- 73 La forme féminine du singulier, *mogłabyś*, est sélectionnée, puisque l'allocutaire a les traits (-masc., +sg.).

6.6. Comparaison des traductions

- 74 Notre proposition : *Czy mogłabyś zrobić reklamę dla mojego forum ?*
- 75 Explication mot à mot de la phrase polonaise : particule interrogative – *pourrais* (2^{ème} personne du singulier du féminin) – *faire* (forme perfective) – *pub* (cas accusatif) – *pour* – *mon* (cas génitif) – *forum* (cas génitif)
- 76 Traduction littérale : *Czy to by Ci przeskadzało zrobić reklamę dla mojego forum ?*
- 77 La deuxième traduction serait comprise par un Polonais à peu près comme : *est-ce que cela (ce fait) te rendrait difficile l'action de faire de la pub ?* La première proposition est une question portant sur la capacité de l'allocutaire (*pourrais-tu faire...*) : nous avons en effet sélectionnée cette stratégie sémantique comme représentative (canonique) de toutes les requêtes du même registre, emploi et degré de politesse. Le sens de la requête est donc conservé dans cette traduction, elle est pragmatiquement et stylistiquement correcte, contrairement à la traduction littérale qui serait proposée par tout traducteur automatique dépourvu de moyens de traitement de PP. Le fait de considérer les PP comme des stéréotypes discursifs propres à chaque langue permet de les traiter comme des suites figées, de déterminer les écarts entre les systèmes de PP dans la langue source et la langue cible et de neutraliser ensuite ces écarts grâce à la définition des points communs aux deux langues, contenus dans la représentation intermédiaire.

Conclusion

- 78 Nous avons montré dans cet article que du fait de la grande diversité des PP, leur transposition mot à mot pouvait donner lieu à des traductions incorrectes, maladroites ou à des contresens. Ces unités méritent donc un traitement particulier en traduction automatique, tout comme les suites non compositionnelles. Notre modèle théorique de traduction des PP tient compte de différents critères pragmatiques, tels que le type d'acte de langage, le registre, les conditions d'emploi, les niveaux de politesse, mais aussi des relations entre les interlocuteurs et, plus généralement, de la diversité formelle et culturelle. Nous avons ainsi cherché à définir ce que les langues peuvent partager comme « dénominateur commun » et à utiliser cette qualité des langues lors de la traduction.

BIBLIOGRAPHIE

- Bourquin G. (1993), « Préalables linguistiques à la construction d'un système de traduction automatique », in Bouillon P., A. Clas (éds), *La traductique*, Montréal : Presses de l'Université de Montréal, pp. 27-34.
- Burger H. (1973), *Idiomatik des Deutschen*, Tübingen : Max Niemeyer.
- Burger H., A. Buhofer, A. Sialm (1982), *Handbuch der Phraseologie*, Berlin, New York.
- Cardey S., P. Greenfield, M. Bioud, A. Dziadkiewicz, K. Kuroda, I. Marcelino, C. Melian, H. Morgadinho, G. Robardet, S. Vienney (2006) "The Classification Sense-Mining System" in *Advances in Natural Language Processing*, Springer-Verlag – LNAI 4139, ISBN 3-540-37334-9, pp. 674-684.
- Dziadkiewicz A. (2007a), *Vers une reconnaissance et une traduction automatique de phraséologismes pragmatiques (application du français vers le polonais)*, thèse de doctorat (confidentielle), Université de Franche-Comté, Besançon.
- Dziadkiewicz A. (2007b), « L'ambiguïté des constructions interrogatives avec *pouvoir* et *vouloir* : peut-on reconnaître formellement les actes d'offre et de requête ? », in Barceló G. J., Lauze A., Patard A. (éds), *De la langue au discours : l'un et le multiple dans les outils grammaticaux*, Actes du IVème Colloque Jeunes Chercheurs Praxiling tenu à Montpellier en mai 2005, Montpellier : Praxiling - Université Paul-Valéry, (sous presse).
- Fonagy I. (1982), *Situation et signification*, Amsterdam : Benjamins.
- Fraser B. (1981), "On apologizing", in Coulmas F. (ed.), *Conversational routine. Explorations in standardized communications situations and prepatterned speech*, The Hague & Paris: Mouton, pp. 260-271.
- Fraser B., E. Rintell, J. Walters (1980), "An approach to conducting research on the acquisition of pragmatic competence in a second language", in Larsen-Freeman D. (ed.), *Discourse analysis in second language research*, Rowley: Newbury House, pp. 75-91.
- Kerbrat-Orecchioni C. (1998), *Les interactions verbales. T.3. Variations culturelles et échanges rituels*, Paris : Armand Colin.
- Kerbrat-Orecchioni C. (2001), *Les actes de langage dans le discours. Théorie et fonctionnement. « Quand dire, c'est faire » : un travail de synthèse sur la pragmatique conversationnelle*, Paris : Nathan.
- Tannen D., P. Comert Oztek (1981), "Health to Our Mouths. Formulaic Expressions in Turkish and Greek", in Coulmas F. (ed.), *Conversational routine. Explorations in standardized communications situations and prepatterned speech*, The Hague & Paris: Mouton, pp. 37-54.
- Wierzbicka A. (2003), *Cross-Cultural Pragmatics. The Semantics of Human Interaction*, Berlin, New York: Mouton de Gruyter.

NOTES

1. Les exemples dans ce paragraphe viennent de Kerbrat-Orecchioni (1998, 2001).
 2. Cependant, nous n'avons pas inclus dans cette étude les constructions purement grammaticales, dépourvues de tout élément lexical fixe, comme par exemple les phrases à l'impératif ou les questions à la 2^{ème} personne du présent, du futur ou du conditionnel (*Ferme la porte ! Tu fais la vaisselle ? Tu vas nettoyer ta chambre ? Tu me prêterais 2 euros ?*).
 3. La même méthodologie a été utilisée dans le système Classificatim, décrit dans Cardey, Greenfield et al. (2006).
 4. A moins d'imaginer des dialogues du genre : - *Comment a-t-il été reçu ?* - *Cordialement*.
 5. Ce module est décrit en détail dans Dziadkiewicz (2007a).
-

RÉSUMÉS

Le terme de phraséologismes pragmatiques (PP) désigne les constructions routinières utilisées conventionnellement dans des situations déterminées. Ces unités, bien que compositionnelles pour la plupart, doivent être traduites non en tant que suites libres, mais en tant que stéréotypes discursifs propres à chaque langue. Nous présentons un modèle théorique de traduction automatique de PP réalisant quelques actes de langage sociaux. Celui-ci est fondé sur l'attribution aux PP d'étiquettes sémantico-pragmatiques composées de critères d'acte de langage, registre de langue, emploi, degré de politesse et stratégie sémantique. Ces étiquettes sont intégrées dans des représentations intermédiaires. Un algorithme de traduction permet ensuite de trouver parmi les PP de la langue cible celui avec la même étiquette (ou avec le plus grand nombre de critères en commun) que le PP à traduire.

The term of pragmatic phraseologisms (PrPs) stands for routine constructions referring conventionally to defined situations. These expressions, even though mostly compositional, have to be translated not as free sequences but as speech stereotypes peculiar to each language. I present a theoretical model of machine translation of PrPs referring to some social speech acts. It is based on assigning to PrPs semantic-pragmatic labels composed of criteria of speech act, language register, use, politeness degree and semantic strategy. These labels are integrated into intermediate representations. Finally I present a translation algorithm which allows finding among the PrPs of the target language the one with the same label (or with the highest number of criteria in common) as the PrP of the source language.

INDEX

Mots-clés : phraséologie, pragmatique, actes de langage, traduction automatique, représentation intermédiaire, stratégies sémantiques, polonais

Keywords : phraseology, pragmatics, speech acts, machine translation, intermediate representation, semantic strategies, polish

AUTEUR

ALEKSANDRA DZIADKIEWICZ

Centre de Recherches en Linguistique Lucien Tesnière, Université de Franche-Comté